



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### Towards an account of epistemic luck for necessary truths

**Citation for published version:**

Collin, J 2018, 'Towards an account of epistemic luck for necessary truths', *Acta Analytica*, vol. 33, no. 4, pp. 483-504. <https://doi.org/10.1007/s12136-018-0360-9>

**Digital Object Identifier (DOI):**

[10.1007/s12136-018-0360-9](https://doi.org/10.1007/s12136-018-0360-9)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

Acta Analytica

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.





# Towards an Account of Epistemic Luck for Necessary Truths

James Henry Collin<sup>1</sup> 

Received: 3 January 2018 / Accepted: 18 June 2018  
© The Author(s) 2018

**Abstract** Modal epistemologists parse modal conditions on knowledge in terms of metaphysical possibilities or *ways the world might have been*. This is problematic. Understanding modal conditions on knowledge this way has made modal epistemology, as currently worked out, unable to account for epistemic luck in the case of necessary truths, and unable to characterise widely discussed issues such as the problem of religious diversity and the perceived epistemological problem with knowledge of abstract objects. Moreover, there is reason to think that this is a congenital defect of orthodox modal epistemology. This way of characterising modal epistemology is however optional. It is shown that one can non-circularly characterise modal conditions on knowledge in terms of epistemic possibilities, or *ways the world might be for the target agent*. Characterising the anti-luck condition in terms of epistemic possibilities removes the impediment to understanding epistemic luck in the case of necessary truths and opens the door to using these conditions to shed new light on some longstanding epistemological problems.

**Keywords** Anti-luck epistemology · Epistemic possibility · Safety condition · Sensitivity condition · Abstract objects

## 1 Modal Epistemology

It is widely held to be a platitude that knowledge excludes luck. To this end, a number of theorists hold that a successful analysis of knowledge requires some modal

---

✉ James Henry Collin  
[james.collin@ed.ac.uk](mailto:james.collin@ed.ac.uk)

<sup>1</sup> University of Edinburgh, Edinburgh, UK

condition, such as safety or sensitivity, that can preserve the sense in which knowledge excludes luck.<sup>1</sup> Safety and sensitivity however have a problem dealing with necessary truths, and little work has been done in addressing the problem. While some efforts have been made to tinker with safety and sensitivity conditions so as to accommodate necessary truths (Sainsbury (1997), Williamson (2000), Weatherson (2004), Becker (2007), Mišćević (2007) Pritchard (2009, 2012), Melchior (2017), Hirvelä (forthcoming)), Roland and Cogburn (2011) make the case, convincingly, that two of these (those of Williamson and Pritchard) are not successful; we will see that *all* of these minor modifications run into problems, and for the same broad reasons. Blome-Tillmann (2017) argues that necessary truths present a serious problem for sensitivity-based accounts of knowledge, and Hales (2016) argues that necessary truths appear to be an intractable problem for modal accounts of luck. Things are not, I think, so bleak, but these problems suggest that a more fundamental kind of change to modal conditions such as safety and sensitivity is required. This paper takes some first steps towards doing this.

Specifically, getting modal anti-luck conditions to work for necessary truths involves transposing them into a different modal key. There are many kinds of modality, but one important variety is *metaphysical possibility*, i.e. *ways the world might have been*. (Following the literature on two-dimensional semantics, I will refer to this as ‘2-possibility’.<sup>2</sup>) Another significant kind of modality concerns *epistemic possibility*, i.e. *ways the world might be*. (Again, following the literature on two-dimensional semantics, I will call this ‘1-possibility’.) Modal epistemologists have always made use of 2-possibility to characterise the modal component of knowledge, whatever they take that to be. But one might think this is something of a historical accident, resulting from the relative neglect, until more recently, of work on 1-possibility.<sup>3</sup> Accident or not, in what follows I make out the case that 1-possibility is better suited than 2-possibility in characterising a modal condition on knowledge. First, I set out the two main rivals in modal epistemology: safety and sensitivity. We will see that both have problems making sense of epistemic luck in cases where the target proposition is non-contingent. And we will see that, worse than this, efforts to tinker with the conditions to avoid the problem all fail. Modal epistemology rests on a mistake, but it is not a mistake fatal to the project. All that is required is a different choice of modality. Once we have set out what 1-possibility is, and have sketched a few details of how to understand 1-possible worlds, we can construct 1-possible counterparts to the orthodox 2-possible safety and sensitivity conditions. As well as removing one impediment to a workable modal epistemology, this will allow the modal component of knowledge to do new explanatory work, for instance in moving forward debates in the philosophy of mathematics and the philosophy of religion where the subject matter is of non-contingent truths.

<sup>1</sup> Note that even *Robust Virtue Epistemologists* such as Zagzebski (1996, 1999), Sosa (2007, 2009) and Greco (2009) who wish to parse knowledge purely in terms of the cognitive abilities of agents must accept some anti-luck condition on knowledge, or otherwise reject the platitude that knowledge excludes luck. In this case, whatever their favoured virtue-theoretic conditions for knowledge, these must at least implicitly contain—i.e. entail—an anti-luck condition. Greco (2012), for instance, develops the case that his recent robustly virtue-theoretic analysis of knowledge can rule out epistemic luck, and in Greco (manuscript) argues that his account entails a safety condition.

<sup>2</sup> See Garcia-Carpintero and Macia (2006).

<sup>3</sup> See, for example, Egan and Weatherson (2011).

## 2 Epistemic Luck: Safety and Sensitivity

The project of anti-luck epistemology takes epistemic luck as being central to our understanding of knowledge.<sup>4</sup> Knowledge excludes luck, but some refinements have to be made; knowledge does not exclude luck *simpliciter*. It may, for instance, be lucky that the proposition in question is true. Perhaps you are a lottery winner. This does not mean that you cannot have knowledge that you are a lottery winner; you may have excellent grounds for your belief. So, *content epistemic luck*—when it is lucky that the proposition is true—is benign. So too is *capacity epistemic luck*—when it is lucky that the agent is capable of knowledge. Perhaps you have made a remarkable recovery from a head injury that ought to have rendered you incapable of thought. You are lucky to be capable of knowledge, but capable of knowledge nonetheless. Another benign form of luck is *evidential epistemic luck*—when it is lucky that you acquire the evidence on which you base your belief. Perhaps you are a detective who discovers, quite by chance, a suspect’s DNA at a crime scene. The element of luck here does not prevent the detective from achieving knowledge. The malign form of epistemic luck is *veritic epistemic luck*—when it is lucky that the agent’s belief is true. Pritchard (2007, 280) offers the following modal condition for veritically lucky true belief:

S’s true belief is lucky *iff* there is a wide class of near-by possible worlds in which S continues to believe the target proposition, and the relevant initial conditions for the formation of that belief are the same as in the actual world, and yet the belief is false.

If knowledge excludes veritic luck, then we require some kind of condition that eliminates veritic luck. Here, safety<sup>5</sup> seems to fit the bill:

S’s true belief *p* is safe *iff* in nearly all (if not all) nearby possible worlds *w* in which S believes *p*, *p* is true.

This needs some refinement. Borrowing an example of Sosa’s (2007, 26), suppose I am hit hard and experience significant pain. On the basis of being hit I believe, and know, that I am in pain. Suppose also that I am a hypochondriac and would have believed myself to be in pain even if I had suffered only a very slight blow. My belief is not safe, as in many of the close worlds in which I hold it, it is false. To avoid this problem, and others like it, a workable safety condition needs to incorporate the basis-for-which/method-in-which/ability-from-which/way-in-which the belief is formed.<sup>6</sup> Pritchard (2007, 2008), for instance, has endorsed this sophisticated and representative version of the safety principle:

*Safety*: S’s belief is safe if and only if in most nearby possible worlds in which S continues to form her belief about the target proposition in the same way as in the

<sup>4</sup> See Pritchard (2005) for a book-length treatment.

<sup>5</sup> Safety is often taken to originate with Sosa (1999) who parsed the condition as ‘a belief by S that *p* [is safe] *iff*: S would believe that *p* only if it were so that *p*’ [142], but a variant of safety appears in Luper (1984).

<sup>6</sup> I will not attempt to adjudicate between these different ways of formulating the matter.

actual world, and in all very close nearby possible worlds in which S continues to form her belief about the target proposition in the same way as in the actual world, the belief continues to be true.

The way in which I actually form my belief involves being, or responding to being, struck hard. Those worlds in which I suffer only a very light blow need not be taken into consideration. If knowledge excludes veritic luck and the amended safety condition gives the correct account of what eliminates veritic luck, then the amended safety condition is a necessary condition for knowledge.

A currently less popular, but important rival anti-luck condition, is the sensitivity principle. Early versions of sensitivity were articulated by Dretske (1971), Goldman (1976) and, most enduringly, Nozick (1981).<sup>7</sup> The principle requires that when an agent believes the proposition that *p* she is sensitive to this fact; in the sense that were that *p* false, she would not believe that *p*. This too needs some refinement. Borrowing an example of Nozick's (1981, 179), 'suppose a grandmother sees that her grandson is well when he comes to visit; but if he were sick or dead, others would tell her he was well to spare her the upset.' The grandmother's belief is not sensitive, since *were* it false that her grandson was well, she would continue to believe it anyway. To avoid this problem, and others like it, a workable sensitivity condition needs to incorporate the basis-for-which/method-in-which/ability-from-which/way-in-which the belief is formed. A sophisticated recent version of the sensitivity condition is endorsed by Becker (2007):

*Sensitivity:* S's belief that *p* is sensitive if and only if, were that *p* false, S would not believe that *p* via the method<sub>n</sub> S actually uses in forming the belief that *p*.<sup>8</sup>

Here, the subjunctive conditional is understood in terms of possible worlds.<sup>9</sup> S's belief that *p* is sensitive if and only if, in the closest possible world<sup>10</sup> in which that *p* is false, S does not believe that *p* via the method<sub>n</sub> S actually uses in forming the belief that *p*. If knowledge excludes veritic luck and the amended sensitivity condition gives the

<sup>7</sup> Though less popular, variants of sensitivity have significant contemporary advocates, like Roush (2005), Black and Murphy (2007b), Murphy (2007a; 2012) and Becker (2007). The sensitivity condition also has particular benefits when it comes to virtue theoretic accounts of knowledge (REFERENCE SUPPRESSED). The reference is: Bolos and Collin (2018).

<sup>8</sup> Some notes on Becker's adoption of the sensitivity principle: Firstly, the subscripted 'n' appended to 'method' indicates that methods are to be individuated narrowly and in a content-specific way. An example of a method<sub>n</sub> is *If what I am looking at now has short legs and floppy ears (and such and so other features), it is a dachshund*. Becker (2007, 106). Note that agents do not need to be aware of the method they are following in order to follow it. Secondly, Becker does not take sensitivity, on its own, to be sufficient to rule out epistemic luck. He combines the sensitivity condition with a modalised and method-relative reliability condition. (ibid.)

<sup>9</sup> The two standard ways of analysing subjunctive conditionals are due to Stalnaker and Lewis. Stalnaker's (1968) method takes a subjunctive conditional  $\phi > \psi$  to hold at a world *w* iff if  $\phi$  is true at some world *x*—and for any world *y* such that  $\phi$  is true, and *x* is at least as close to *w* as *y*—then  $\psi$  is true at *x*. Or, more demotically,  $\phi > \psi$  is true iff in the closest possible world in which  $\phi$  is true,  $\psi$  is also true. Lewis' (1973) method takes  $\phi > \psi$  to hold iff either  $\phi$  is true at no worlds, or there is a world *x* in which  $\phi$  is true, and for all worlds *y* if *y* is at least as similar to the actual world as *x* then  $\phi \rightarrow \psi$  is true at *y*. To make the analysis work, Nozick (1981, 174) interpreted subjunctive conditions in a stronger way:  $\phi > \psi$  is not automatically true when  $\phi$  and  $\psi$  are both true at the actual world, but only when  $\psi$  is true at *nearby*  $\phi$  worlds.

<sup>10</sup> Or *worlds* (plural). The sensitivity theorist has a choice here about which formulation to adopt. See Becker (2007 ch.2, B) for the differences this makes.

correct account of what eliminates veritic luck, then the amended sensitivity condition is a necessary condition for knowledge.

## 2.1 The Problem

These modal conditions, when incorporated into an account of knowledge, are designed to prevent that account from over-generating: counting as knowledge cases that are epistemically lucky in a way that is incompatible with knowing. However, they seem incapable of doing so when the propositions in question are necessarily true. A quick glance at possible world semantics exposes the problem. Consider the following clause:

$$(\text{Nec}) \ v(\Box p, w) = T \text{ iff for every world } w' \text{ in } W, \text{ such that } Rww', \ v(p, w') = T$$

where  $v(p, w)$  is the truth value of  $p$  at world  $w$ ,  $W$  is the set of possible worlds, and  $R$  is an accessibility relation on worlds. Intuitively, this states that  $\Box p$  is true at a world  $w$  when  $p$  is true at all possible worlds (accessible from  $w$ ). From (Nec) it follows that:

(Nec-Safe) If  $\Box p$ , then in all nearby possible worlds  $w$  in which  $S$  forms a belief  $p$  in the same way as in the actual world,  $p$  is true.

and

(Nec-Sensitive) If  $\Box p$ , then in the closest possible world in which  $\neg p$ ,  $S$  fails to belief  $p$  via the same method as in the actual world.

In other words, if  $p$  is necessarily true, then  $S$ 's belief  $p$  is automatically safe and sensitive. But surely there are cases of luckily true belief in necessary propositions. For instance:

*LUCKY 8-BALL:* Gullible Joe forms beliefs by shaking a lucky 8-ball, which, instead of predicting the future, only states necessary a posteriori identity claims. Gullible Joe knows nothing about the provenance of the 8-ball, and believes whatever the 8-ball tells him.

Gullible Joe's beliefs are necessarily true, and, as such, trivially safe. Yet, Gullible Joe does not have knowledge of these claims. Moreover, it is natural here to say that Joe does not have knowledge as a result of his belief being only luckily true. So, there are cases where the safety condition does not achieve its end of eliminating epistemic luck. Joe's belief is lucky, but the world modally over-cooperates and makes his belief safe.<sup>11</sup>

At this juncture, one option is simply to restrict our account to contingent propositions, but even this is not entirely sufficient. Suppose that, as a result of taking a powerful hallucinogenic, Gullible Joe forms the belief that objects contract when in motion. There are no nearby possible worlds in which Lorentz Contraction does not

<sup>11</sup> See Roland and Cogburn (2011) for a different presentation of this objection, and a convincing case that the attempts of Williamson (2000) and Pritchard (2009) to overcome it are not successful.

take place—as such worlds are nomologically impossible—thus Joe’s belief is safe. So, if we are to restrict the account, then we must go further and restrict it to ‘fully contingent’ propositions: propositions which are nomologically as well as metaphysically or logically possible. This option is not appealing; what we want is an account of epistemic luck that is applicable to all propositions. Joe’s belief here seems luckily true; a universally applicable account of epistemic luck could shed light on what is common to all cases of luckily true belief. On the other hand, an account of epistemic luck that does not tell us what is common to all cases of luckily true belief—even if materially adequate over a restricted range of cases—could never constitute an analysis of epistemic luck, or even a set of necessary conditions for luck, since it could never make explicit what, in general, makes the difference between a belief exhibiting and not exhibiting epistemic luck.

## 2.2 Attempted Solutions

I have claimed that we need to modify or extend our account of epistemic luck in some way in order to accommodate luck pertaining to necessary truths. However, there is already an account of how this might be achieved:

[A]ll we need to do is to talk of the doxastic result of the target belief-forming process, whatever that might be, and not focus solely on belief in the target proposition. For example, if one forms one’s belief that  $2 + 2 = 4$  by tossing a coin, then while there are no near-by possible worlds where that belief is false, there is a wide class of near-by possible worlds where that belief-forming process brings about a doxastic result which is false (e.g., a possible world in which one in this way forms the belief that  $2 + 2 = 5$ ). The focus on fully contingent propositions is thus simply a way of simplifying the account; it does not represent an admission that the account only applies to a restricted class of propositions. (Pritchard (2009, 3))<sup>12</sup>

It may be thought that a deficiency of this approach is that it provides a disjunctive account of epistemic luck, treating contingent or fully contingent propositions one way and non-contingent or not fully contingent propositions another way. If there is something common to all cases of luckily true belief, only a unified account of epistemic luck could capture this. But perhaps veritic epistemic luck isn’t a unified phenomena after all, and distinct analyses of epistemic luck are required for contingent

<sup>12</sup> See also Pritchard (2012). This approach is taken from Mišćević (2007) where he suggests a requirement of *Agent Stability* for necessarily true propositions: ‘If an agent knows a priori a (necessary) proposition  $p$ , then, in most nearby possible worlds in which she forms her belief about  $p$  in a slightly different way or with slightly changed cognitive apparatus as in the actual world, that agent will also come to believe that  $p$ .’ [ibid.: 62] In fact, this kind of response is pre-empted by Sainsbury, who states ‘It is easily possible for me to be wrong in believing that  $p$  (even if it is true that  $p$ ) iff at some world “close” to the actual world the actual episode of forming the belief that  $p$  (or the counterpart for this episode) is one in which a false belief is formed. In some cases, this is because the same belief is, but is false at the close alternative world. In other cases, it is because a different, and false, belief is formed at the close world’ Sainsbury (1997, 913). Williamson also claims that safety-based accounts of knowledge can handle necessary truths, since in cases of luckily true belief in a necessary truth the ‘method by which [the belief is] reached could just as easily have led to a false belief in a different proposition’ Williamson (2000, 182).



and necessary propositions. This is implausible. The lucky 8-ball above stated necessary truths, but if Gullible Joe formed true beliefs on the basis of a lucky 8-ball that stated contingent truths his beliefs would also be epistemically lucky, *and in the same way*. As well as being implausible, it's irrelevant. We would still want an analysis of epistemic luck for necessary truths, and would still be lacking one. However, the criterion could be formulated so as to handle both kinds of cases. Weatherson contrasts *Content-safety* where '*B* is safe iff *p* is true in all similar worlds' and *Belief-safety* where '*B* is safe iff *B* is true in all similar worlds' (Weatherson (2004, 378)). Incorporating the notion of a basis to avoid the kind of counterexamples discussed above, we might endorse:

(Safety\*) S's belief is safe iff in all nearby possible worlds in which S forms a belief on the same basis as in the actual world S forms a true belief.

Safety\* provides a unified account of luck in both cases where the proposition believed is contingent and where it is necessary. A more pressing problem though—both for the disjunctive account, and for safety\*—is that they rely on the 'doxastic result of the target belief-forming process' being somehow unstable; but this need not be the case. There is nothing to preclude examples where (i) what is believed is necessary, (ii) the proposition believed is fixed across close possible worlds, and (iii) the belief is luckily true. (i–iii) all obtain in the lucky 8-Ball case, for example.

I have criticised safety and sensitivity conditions on the grounds that there are cases in which these conditions hold for an agent, yet the agent fails to possess knowledge. This may seem unfair, since both safety and sensitivity are proffered as necessary rather than sufficient conditions for knowledge. However, the claim here is not that these modal conditions are incorrect, but instead that we ought to seek out a well-motivated modal condition which is superior, which captures more of the contours of knowledge and epistemic luck. Coming upon necessary conditions for knowledge is no great feat; the trick is to find the most *informative* necessary conditions possible. It is a necessary condition for knowing *p* that one does not come to believe *p* on the basis of counter-reasons; but although this is a necessary condition for knowledge, it is not a particularly informative one. The stronger a necessary condition is, the more informative it is. In the same way, the examples above show that safety is not *sufficient* for anti-luck and so is not an anti-luck condition. If an alternative could be found which did not face these counterexamples, it would be closer to a sufficient anti-luck condition and so more explanatory.

Of course, we may hope to explain the failure to gain knowledge by appealing to some other principle. In this case, reliability will not do the trick, as Joe's belief-forming processes reliably, indeed invariably, lead him to these necessarily true beliefs. Becker (2007, 101–4) attempts to deal with cases of lucky truth in necessary propositions by claiming that the methods used in such cases will not, in general, be *reliable*. Becker contrasts a *careless* maths student who, on a whim, adopts a (fortuitously) sound algorithm for solving problems and a *naturally gifted* maths student who always adopts sound algorithms for solving problems. 'The [method] used by the careless math student seems to be fleeting, whereas those used by the naturally gifted math student are not.' (ibid. p.102). Fleeting methods are not, in general, reliable. However, it should



be clear that the 8-ball method, although epistemically lucky, is *not* fleeting in the relevant way, and, as such, there are cases that cannot be dealt with in the manner Becker suggests.

Nor will appeals to justification or abilities help in all cases. We need not suppose that Joe has failed to exercise his epistemic abilities, that he is epistemically blameworthy or irrational in any way, or that he lacks justification in some other sense; he may be the victim of a widespread and sophisticated collusion. For example:

*PARENTHOOD TEST:* Discerning Joe provides a DNA sample to a machine which tells its users who their parents are. He is told that his parents are cousins. For everyone other than Joe, the machine goes through a 100% accurate checking procedure to provide the parenthood result. The machine however has been programmed to tell Discerning Joe that his parents are cousins without checking whether this is the case. Joe's parents are in fact cousins.

If, with Kripke, we take it that people have their parents essentially, then Discerning Joe's belief is necessarily true and so trivially safe and sensitive. But even if there *were* some condition which acted as a band-aid for cases such as these, to invoke it would miss the point. Joe is *lucky* that his beliefs are true; an account of epistemic luck should try to capture this. One might hope that we can bolster the notion of ability in a way that will avoid the problem. For instance, Sosa (2007) understands knowledge as *apt performance*<sup>13</sup>:

[W]e can distinguish between a belief's accuracy, i.e. its truth; its adroitness, i.e. its manifesting epistemic virtue or competence; and its aptness i.e., its being true because competent (ibid, 23).

A belief counts as knowledge so long as it manifests these three as: it is accurate, adroit—arriving at the belief is the result of manifesting an epistemic virtue or competence—and apt—that the agent has arrived at a true belief is because of or a result of that manifestation of competence. This however, still falls foul of PARENTHOOD TEST. Let's stipulate that, in this scenario, operating the machine requires skill on the part of the agent. Let's stipulate further, that if Joe lacked this skill, and did not use the machine correctly, then it would have provided an *incorrect* result. In this case, Joe's belief is accurate: the machine correctly tells Joe that his parents are cousins. It is also adroit: Joe manifests epistemic competence in using the machine correctly; the same epistemic competence that others manifest when they gain *knowledge* from the machine. Finally, Joe's belief is apt: he gains a true belief because of his competence. Had he not been a competent operator of the machine, his belief would have been false.<sup>14</sup> Yet, Joe still lacks knowledge, because his belief is luckily true. Something else is needed.

<sup>13</sup> Melchior (2017) appeals to the idea of an apt performance to account for epistemic luck in cases of necessary truths.

<sup>14</sup> Of course the machine itself also played a role in Joe arriving at a true belief. But if we take this to reduce the role of Joe's competence sufficiently to undermine the aptness of his belief, then we would have to say the same thing about other agents who use the machine. On Sosa's account, however, that would mean (implausibly) that *nobody* could gain knowledge through the machine.

In a similar vein to Melchior (2017), and drawing on Sosa (2010), Hirvelä (forthcoming) defends the following safety condition:

GLOBAL SAFETY:

S's belief that *p*, which belongs to her subject matter of inquiry *Q*, is safe if and only if:

- (i) in most possible worlds, and in all of the very closest possible worlds, where *S* believes in a proposition that belongs to *Q* via the same virtuous method *V* that *S* uses in the actual world, *S*'s belief is true.

Here, virtuous methods are relative to circumstances, and themselves defined in terms of their propensity to produce true beliefs:

VIRTUE:

A subject *S*'s belief that *p*, which belongs to a field of propositions *F*, is virtuously formed via method *V*, in circumstances *C* is and only if:

- (i) *S* has an inner disposition *D* to attain correct doxastic attitudes with respect to propositions that belong to *F*,
- (ii) *S* is in *C* and believes that *p*
- (iii) the fact that *S* believes that *p*, via *V* is due to exercising *D*

The field of propositions *F* is 'restricted in terms of the agent's subject matter of inquiry' (ibid.). The details of how this restriction is to work are not relevant here. The trouble is that, as with the triple-A account of knowledge, PARENTHOOD TEST clearly meets all the criteria required for GLOBAL SAFETY. Discerning Joe has an inner disposition to attain correct doxastic attitudes with respect to propositions concerning who his parents are, believes that his parents are cousins in the relevant circumstances, and believes this as a result of exercising his inner disposition. As such, Joe has a virtuous method. Moreover, in most possible worlds and in all very close possible worlds where Joe believes that his parents are cousins (where the subject matter of inquiry has to do with who Joe's parents are), Joe believes a proposition that belongs to this subject matter via the virtuous method he uses in the actual world, and that belief is true. Yet, Discerning Joe's belief is luckily true.

In fact, there is reason to think that *no* permutation of safety or sensitivity in terms of possible worlds could capture the contours of epistemic luck. By dealing only with possible worlds, safety and sensitivity build in a structural assumption that only possibly true propositions are epistemically relevant to agents. But we often find ourselves in situations where propositions that are metaphysically impossible are on the table, for instance if we are trying to work out who our biological parents are, from a group of (epistemically) possible candidates. Just as the world can be such that (luckily) it makes our beliefs true—as in the Gettier cases—and just as the world can be

such that (luckily) it makes our beliefs reliable, so too the world can be such that (luckily) it makes our beliefs safe or sensitive. In each instance, features of the world, irrelevant to any cognitive activity on the part of the agent, collude to make her belief true.

### 3 Ways the World Might Be

We have, up until this point been dealing with 2-possibility, but 1-possibility may be of more use to us. 1-possibility is most commonly invoked in discussions of *epistemic possibility*: ways the world might be. 1-possibility has in the first instance to do with ways the *actual* world might turn out to be for a person, given that person's epistemic predicament. 2-possibility deals with *alternatives* to the actual world: ways in which our universe could have been different, regardless of anyone's epistemic predicament. For all I, or anyone, knows, it might be that  $P = NP$  or that Goldbach's conjecture can be proven. For all I know infallibly, it might be that my biological parents are not identical with the people I think are my biological parents, that water is not  $H_2O$ , or that classical logic is unsound. These are all 1-possibilities for me, but they may not be 2-possibilities. If it actually is not the case that  $P = NP$  or that Goldbach's conjecture can be proven, or that my biological parents are not identical with the people I think are my biological parents, or that water is not  $H_2O$ , or that classical logic is unsound, then it is metaphysically necessary that this is so. If these claims are actually false, then they are metaphysically impossible, but they remain epistemically possible for me.

One characteristic feature of epistemic possibility then is that what is epistemically possible may not be metaphysically possible. To avoid confusion, one ought not to talk of *possible worlds* when one deals with epistemic possibility as, on the usual understanding, there are no possible worlds in which, say, Hesperus is not Phosphorus, yet it is epistemically possible for many agents that Hesperus is not Phosphorus. As such, we talk of *scenarios* which *verify* or *falsify* sentences or propositions: a scenario verifies a sentence when the sentence is true in that scenario, and falsifies a sentence when the sentence is false in that scenario.

I will largely adopt Chalmers' (2011) account of epistemic possibility. Clearly, one cannot straightforwardly identify scenarios with possible worlds for the reason just mentioned: some scenarios are metaphysically impossible. One ostensible option here (which I'll rule out in a moment) is to attempt to evade the problem, still using the apparatus of metaphysically possible worlds, by distinguishing between *verification* and *satisfaction*, and by making scenarios 'centred' worlds, where a centred world is an ordered triple  $\langle w, \varphi_1, \varphi_2 \rangle$  of a (metaphysically possible) world, an individual and a time. In this way, we augment our third-person description of a world with indexical information regarding the world's centre. If an individual  $x$  is at the centre of a world  $w$ , then we can make use of a predicate  $\varphi$  which is only true of  $x$  at  $w$ . Our ordered triple then will be a complete objective description of a world, along with sentences of the form 'I am  $\varphi_1$ ' and 'Now is  $\varphi_2$ '. Even though no metaphysically possible world *satisfies* 'Water is XYZ'—which is to say, there are no scenarios where it is *true* that water is XYZ—we can now say (roughly) that a scenario *verifies* 'Water is XYZ' if it is a centred world where 'water' picks out the watery stuff with which the individual at the centre of the world is familiar, which, at this world, is XYZ. Issues still remain, a particularly salient one being 'strong necessities': necessary truths that

are verified by all centred worlds. If God exists necessarily, for instance, then there will be no centred worlds where God does not exist. Yet, it is epistemically possible that God does not exist. Some too equate nomological necessity with metaphysical necessity. If that view is correct then, there will be no centred worlds which exemplify different physical laws from our own; yet, it is epistemically possible that, for instance, Lorentz contraction or time dilation do not take place.

A better strategy is to bypass such thorny issues altogether and construct epistemic scenarios from the ground up, independent of possible worlds. For even if we concur with those, like Chalmers (2002), who deny that there are any strong necessities, epistemic possibility is orthogonal to such metaphysical issues, and our account of epistemic possibility should reflect this autonomy. Instead then of identifying scenarios with triples of the form  $\langle w, \varphi_1, \varphi_2 \rangle$ , as we considered above, we can think of scenarios as sentence types of an ideal language  $L$ .<sup>15</sup> A claim is verified by a scenario just in case the sentence type  $d$  that specifies the scenario says that it is the case. Intuitively, a scenario is a description of a way things might be, and a claim is verified at a scenario if and only if this description says things are as the claim says things are. If a sentence says 'Water is XYZ', then that sentence is verified at all and only the scenarios that describe water as being XYZ. There is nothing more to the proposition that  $p$  being 'true in' a scenario than that scenario saying that  $p$  is true.  $L$  will permit infinite sentences, in order to describe some kinds of infinite scenarios. It must also use only 'epistemically invariant' expressions: expressions whose epistemic import will not change from context to context, or utterance to utterance. It is often thought, for obvious reasons, that names and natural kind terms are not epistemically invariant (apart from perhaps in the mouth of God), nor are context-sensitive terms. A *fully specified* scenario can be identified with a sentence  $d$  of  $L$  which is epistemically possible and for which there is no sentence  $s$  such that  $d \& s$  and  $d \& \neg s$  are both epistemically possible. Intuitively, this is just to say that, for any given claim, a fully specified scenario takes a stand on whether it is true or false. Fully specified scenarios are comprehensive in the way that metaphysically possible worlds are comprehensive.

A maximally liberal conception of epistemic possibility would treat all sentences of  $L$  as epistemically possible. If a subject knew nothing whatsoever (if such an idea is coherent), all scenarios would be epistemically possible for that subject. This is a kind of epistemic possibility that captures *what might be the case prior to what anyone knows*. Call this *very deep epistemic possibility*. How we choose to carve up epistemic space depends on our purposes, and it may be that there are uses for very deep epistemic possibility. However, the conception is too liberal for the case at hand. We are interested in what is epistemically lucky for agents for whom not everything is an epistemic possibility in the relevant sense. A less liberal conception of epistemic possibility treats as epistemically possible all and only those sentences that are compatible with what the relevant agent knows. Call this *strict epistemic possibility*. Since I know through the testimony of experts that water is H<sub>2</sub>O, it is strictly epistemically impossible, for me, that water is not H<sub>2</sub>O. This conception of epistemic possibility however is too *il*liberal for the case at hand. An analysis of knowledge in terms of epistemic possibility, which in turn is defined in terms of what one knows, would be circular. The kind of epistemic possibility we are interested in lies somewhere in between very deep epistemic possibility and strict epistemic possibility. One option here is the

<sup>15</sup> Again, I am following Chalmers (2011) here.

idealised version of epistemic possibility Chalmers (2011) works with: call this *deep epistemic possibility*. Here, a sentence  $s$  is epistemically necessary just in case it is knowable a priori, in an idealised sense that ‘abstracts away from contingent cognitive limitations’ (Chalmers 2011, 68). On this understanding, ‘If there is any possible mental life that starts from a thought and leads to an a priori justified acceptance of that thought, the thought is a priori’ (ibid.) in the relevant idealised sense. Anything that could, in principle, be ruled out a priori is deeply epistemically impossible. Deep epistemic possibility is useful for modelling the knowledge of idealised agents, and is closer to what we need, but since the agents we are concerned with aren’t ideal, it is still too illiberal for us. If an agent has not *yet* ruled out a priori that  $56 \times 83 \neq 4642$ , then that  $56 \times 83 = 4642$  remains an epistemic possibility for that agent. If the agent arrives at the true belief that  $56 \times 83 = 4648$  by tossing a coin, the agent’s belief will still be luckily true. However, the idealised notion of deep epistemic possibility rules this out, because, on according to that notion, there are no epistemically possible worlds in which  $56 \times 83 = 4642$ . What we should say then is this: the relevant sense of epistemic possibility has to do with what *very deep* epistemic possibilities the agent has *in fact* ruled out. A scenario  $w$  is epistemically possible (in the relevant sense) for an agent  $S$ , if  $S$  has not ruled out  $w$  a priori. Call this relevant sense of epistemic possibility ‘1-possibility’. We can then say:

A proposition  $p$  is 1-possible for an agent  $S$  iff  $S$  has not ruled out  $p$  a priori.

and

A scenario  $w$  is 1-possible for an agent  $S$  iff  $S$  has not ruled out  $w$  a priori.

A final analysis of knowledge, then, would come in two steps. The first step would involve analysing a posteriori knowledge in terms of *ruling out* a priori, and *ruling out* a priori would be analysed in terms of something else. This would be neither troubling nor surprising. The a priori is very different from the a posteriori and raises its own set of issues. We might expect an account of a priori knowledge to be very different from an account of a posteriori knowledge. The goal here is not to arrive at the final analysis of knowledge, but only to work towards a more satisfactory account of epistemic luck. However, we can say something instructive about the various ways the former might go.

Some have argued that there is no such thing as a priori knowledge.<sup>16</sup> If this is the case, then things are more straightforward. 1-possibility would here be equated with *very deep epistemic possibility*, and the analysis of knowledge would not require two steps. In this case, all scenarios would be, in some sense, relevant to whether the agent’s belief was luckily true, as no scenarios would be ruled out a priori. In fact, this will make no difference to the situations we consider. As we will see, what matters for safety is that in most *nearby* scenarios and all very close nearby scenarios, in which the agent continues to form her belief in the same way as the actual world, the belief continues to be true. Whether the belief is safe will depend on the belief-forming method the agent is using. Similarly, what matters for sensitivity is that in the closest scenario(s) in which the target belief is false,  $S$  does not continue to hold that belief via the method  $S$  actually uses in forming that belief. Again, whether the belief is safe will

<sup>16</sup> See, for instance, Devitt (2014).

depend on the belief-forming method of the agent. There may be some situations in which safety or sensitivity is more easily achieved if there is such a thing as a priori knowledge. If there is no a priori knowledge, there are more epistemically possible scenarios than if there is such a thing as a priori knowledge. There may be cases in which some of these scenarios become relevant to the safety or sensitivity of someone's belief, and, in fact, in such a way as to undermine the safety or sensitivity of those beliefs. This, though, is exactly what we should expect, and would be modelled by the safety and sensitivity conditions that will be put forward shortly.

What if there *is* such a thing as the a priori? One might worry here that this would make the account circular. We would be analysing a posteriori knowledge in terms of a priori knowledge. This however is too quick. Because the a priori and the a posteriori are so different, there is no reason to suppose at the outset that any circularity would be involved. For the analysis to be circular, it would have to be the case both that a posteriori knowledge was analysed in terms of a priori knowledge, *and* that a priori knowledge was, in turn, analysed in terms of a posteriori knowledge. Take a broadly Carnapian notion of the *relative* a priori, but given a more Hilbertian axiomatic treatment to avoid the problems that beset Carnap's own accounts (see, e.g. Awodey 2007).<sup>17</sup> If one understands the a priori in broadly Carnapian terms, then the a priori will be analysed in terms of linguistic competence. Recognising, and hence knowing, that something is a priori, or ruled out a priori, is a matter of understanding the language one has chosen to adopt. Given this approach to the a priori, there is no apparent danger of circularity. But what of other understandings of a priori knowledge? How to understand the a priori is a vexed issue, and a variety of options are on the table. Whether there would be some kind of circularity here would depend on how a priori knowledge was eventually understood. As I mentioned, there is no particular reason to think from the outset (and even some reason to doubt) that a priori knowledge would end up being analysed in terms of a posteriori knowledge. But take the 'worst case scenario': that your favourite account of a priori knowledge *did* result in a circular analysis. In fact, even this 'worst case scenario' is not so very bad. A circular necessary condition for something can still be *informative*; it can still shed important light on the concept being explicated. Williamson—who famously holds that knowledge cannot be analysed—makes this point with respect to the safety condition, which he endorses:

The obvious worry about such a circular account is that it is uninformative. But circularity does not entail uninformativeness, as Nelson Goodman pointed out long ago. When David Lewis gave the semantics of counterfactual conditionals in terms of similarity relations between possible worlds, his methodology was to work out what respects of similarity carry most weight from which counterfactuals are true. He readily conceded that his statement of the truth-conditions for counterfactuals is vague, but insisted that what matters is that its vagueness matches the vagueness of the original (Lewis 1986: 91–5). In many tricky examples, Lewis's account does not deliver a clear independent prediction as to the truth-value of a counterfactual conditional. Nevertheless, his account is highly informative, especially about structural matters such as the logic of

<sup>17</sup> This is the understanding of the a priori I take to be the most plausible, but nothing I argue for here rests on this.

counterfactuals. Likewise, the role of the safety account in *Knowledge and its Limits* is not to deliver clear independent predictions as to the truth-values of knowledge claims in particular tricky examples. Nevertheless, it is highly informative, especially about structural matters (Williamson 2009, 305-6).

Even if (as I suspect is *not* the case) the final ‘analysis’ of knowledge ended up being circular in this way, this is no reason to throw it out. On the contrary, unpacking necessary conditions for knowledge is still informative, as it allows us to investigate the structural features of knowledge, even when this does not end up providing a non-circular analysis of knowledge. The point here is that traditional safety and sensitivity conditions *do not* correctly characterise the structural features of knowledge, and need to be replaced by conditions that can.

Having given this sketch of epistemic possibility,<sup>18</sup> its relevance should be clear. That the metaphysically impossible is often epistemically possible means that epistemic possibility may be better placed as a tool for analysing epistemic luck, as it may help us provide an account of veritic luck regarding belief in metaphysically necessary propositions. As the safety condition traditionally made use of 2-possibility, we can refer to it as ‘safety<sub>2</sub>’. Mirroring Pritchard’s earlier formulation of the principle, we can simply replace safety<sub>2</sub> with safety<sub>1</sub>:

*Safety<sub>1</sub>*:

S’s belief is safe if and only if in most nearby scenarios  $w \in W_S$  in which S continues to form her belief about the target proposition in the same way as the actual world, and in all very close nearby scenarios  $w \in W_S$  in which S continues to form her belief about the target proposition in the same way as the actual world, the belief continues to be true.

where  $W_S$  is the set of scenarios that are 1-possible for S. This new formulation has a number of advantages. For one thing, it allows us to see how veritic luck can infect belief in necessary propositions: if a proposition  $p$  happens to be metaphysically necessary, S’s belief  $p$  isn’t automatically trivially safe.

### 3.1 Refinements

However, more work needs to be done; refinements to the safety<sub>1</sub> principle are required. It seems that some epistemically possible scenarios—scenarios which are nomologically impossible or which involve massive changes in particular fact—are still too far away for lucky beliefs to be unsafe. Consider an agent who forms a luckily true belief that water is H<sub>2</sub>O. Are there really close possible scenarios in which water is not H<sub>2</sub>O? Scenarios in which the liquid that occupies 71% of the earth’s surface is different to that of the actual world would be quite distant, or, at any rate, too far away to be troubled by a safety requirement. As such, the account needs to be finessed.

<sup>18</sup> As mentioned, for details see Chalmers (2006, 2011), from which the preceding discussion draws.



One option is to adopt a different kind of modal requirement. Mirroring Becker's earlier formulation this time, we can simply replace Sensitivity<sub>2</sub> with Sensitivity<sub>1</sub>:

Sensitivity<sub>1</sub>:

S's belief that  $p$  is sensitive if and only if, in the closest scenario(s)  $w \in W_S$  in which that  $p$  is false, S does not believe that  $p$  via the method <sub>$n$</sub>  S actually uses in forming the belief that  $p$ .

Using sensitivity<sub>1</sub> as an anti-luck condition deals neatly with the above example. Those who prefer safety conditions also have a way of dealing with these cases, although it involves a bit more work.<sup>19</sup> 1-possible worlds are fully specified scenarios. Recall that a fully specified scenario is identified with a sentence  $d$  of  $\mathcal{L}$  which is epistemically possible and for which there is no sentence  $s$  such that  $d \& s$  and  $d \& \neg s$  are both epistemically possible. But it is also possible to deal in *partial* scenarios: scenarios which are not fully specified. Instead of ordering scenarios relative to the actual world—which can be thought of as a fully specified scenario—we can order scenarios relative to a partially specified scenario  $w'$  which verifies only the set of sentences that specifies the way the agent forms the belief at the actual world, leaving open other aspects of the world. Call this set  $\{\psi : \psi \in \Sigma(S, @)\}$  (the set of sentences  $\psi$  such that  $\psi$  is in the set of sentences  $\Sigma$  specifying person S's belief-forming method at the actual world @). In cases where an agent's belief that water is H<sub>2</sub>O is based on the roll of a die, for instance,  $w'$  would be 'agnostic' with regard to the chemical composition of water; hence, there would be many nearby scenarios in which water is not H<sub>2</sub>O. In this way, we could reformulate the safety<sub>1</sub> condition as:

Revised Safety<sub>1</sub>:

S's belief is safe if and only if in most nearby\* scenarios  $w \in W_S$  in which S continues to form her belief about the target proposition in the same way as the actual world, and in all very close nearby scenarios  $w \in W_S$  in which S continues to form her belief about the target proposition in the same way as the actual world, the belief continues to be true.

where 'nearby\*' means nearby to the partial scenario  $w'$ , which verifies  $\{\psi : \psi \in \Sigma(S, @)\}$ . This removes the in-principle barrier to treating cases where the target proposition is only false in distant metaphysically possible worlds.

With Sensitivity<sub>1</sub> and Revised Safety<sub>1</sub> (henceforth just 'Safety<sub>1</sub>') in hand, we can see how they handle the problem cases. First though, note that they handle the standard sorts of cases that motivate modal conditions in the first place.

<sup>19</sup> In fact, making use of sensitivity<sub>1</sub> is my preferred approach, owing in part to the sensitivity principle's deep theoretical unity with virtue-theoretic approaches to knowledge (reference suppressed The reference is: Bolos and Collin (2018)). Nothing here hinges on this preference however.

*The Stopped Clock*<sup>20</sup>: At midday, S looks at a clock with both hands pointing up and forms the justified true belief that it is midday on that basis. Unbeknownst to S, the clock stopped at exactly midnight last night. S lacks knowledge because S's belief is veritically lucky. S has not determined the time a priori, so there are 1-possible scenarios in which it is not midday. Take Sensitivity<sub>1</sub> first. In the closest scenario(s) in which it is not midday (i.e. when S looks at the clock at a slightly different time), S continues to believe that it is midday via the method S actually uses in forming the belief that it is midday. S's belief is not sensitive<sub>1</sub>: the right result. Now take Safety<sub>1</sub>. There are many scenarios nearby (to the partial scenario which verifies the way S forms her belief in the actual world) in which it is not midday, but in which S continues to form the same, but here false, belief about the target proposition in the same way as the actual world. These are the nearby scenarios in which S looks at the clock at a slightly different time. S's belief is not safe<sub>1</sub>: the right result.

Barn Façade County<sup>21</sup>: Unbeknownst to S, S is driving through Barn Façade County in which almost all the barns are made to look like actual barns. S looks at the sole barn in the county that is a real barn, and thereby forms the true belief that he has seen a real barn. S has not determined that he has seen a barn a priori, so there are 1-possible scenarios in which he has not seen a barn. Take Sensitivity<sub>1</sub> first. In the closest scenario(s) in which S is not looking at a barn (i.e. when S looks at one of the many surrounding fake barns), S continues to believe that he is looking at a barn via the method S actually uses in forming the belief that he is looking at a barn. S's belief is not sensitive<sub>1</sub>: the right result. Now take Safety<sub>1</sub>. There are many scenarios nearby (to the partial scenario which verifies the way S forms his belief in the actual world) in which S is not looking at a real barn, but in which S continues to form the same, but here false, belief about the target proposition in the same way as the actual world. These are the nearby scenarios in which S looks at one of the many surrounding fake barns. S's belief is not safe<sub>1</sub>: the right result.

We can also see how this plays out with respect to the lucky 8-ball case. Recall the example:

*LUCKY 8-BALL*: Gullible Joe forms beliefs by shaking a lucky 8-ball, which, instead of predicting the future, only states necessary a posteriori identity claims. Gullible Joe knows nothing about the provenance of the 8-ball, and believes whatever the 8-ball tells him.

Take Sensitivity<sub>1</sub> first. For the sake of specificity, let's say that Joe's belief-forming method can be parsed as something like *For any claim, if I read that claim on a lucky 8-ball, that claim is true*, and that the lucky 8-ball was not designed to be accurate.<sup>22</sup> Joe's belief is not sensitive<sub>1</sub>, as it does not track the facts across relevant epistemically possible scenarios. In the closest scenario in which any of the claims stated on the 8-ball do not obtain, Joe continues to believe them. The closest scenario in which any necessary truth is false may be 'distant' from the actual world, but when making use of a sensitivity condition this is neither here nor there. We simply look to the *closest* scenario in which the target belief is false ('closest' does not entail 'close'). Because, by stipulation, the contents of the 8-ball aren't

<sup>20</sup> See Russell (1948, 170)

<sup>21</sup> See Goldman (1976).

<sup>22</sup> This specificity is not trivial. If someone knew that a given lucky 8-ball was designed to be accurate and made by people with the relevant knowledge—say as some kind of mnemonic tool—then they could come to have knowledge by reading the 8-ball.

hooked up to the facts in any relevant way, the closest scenario in which the claim Joe is actually looking at is false is one in which the 8-ball is the same as the actual world. In this scenario, Joe forms a false belief. As such, his belief is not sensitive<sub>1</sub>.

Safety<sub>1</sub> is less straightforward, but it is still easy to see how the example pans out. In the actual world, of course, all the necessary truths stated on the lucky 8-ball obtain; this is why stable beliefs about them are trivially safe<sub>2</sub> and sensitive<sub>2</sub>. But Joe's belief-forming method is irresponsibly narrow. The basis for his target belief is simply that he has read the claim on an 8-ball. So Joe's basis for belief includes facts about the 8-ball, but doesn't include any facts about the relevant necessary truths, or, for that matter, anything that might be connected to those truths. The partial scenario specifying Joe's belief-forming method leaves open (says nothing about) whether the necessary a posteriori identity statements obtain. As such, the scenarios in which these identity statements obtain and the scenarios in which these identity statements don't obtain are equally nearby\*. This is enough to see that, given Joe's belief-forming method, the facts could 'go either way'. The possible scenarios in which Joe's belief is false are more numerous and just as close as the possible scenarios in which Joe's belief is true. Say that the necessary a posteriori identity statement Joe reads is 'Water = H<sub>2</sub>O.' The partial scenario which verifies Joe's belief-forming method at the actual world takes no stance on the identity of water and H<sub>2</sub>O. As a result, scenarios in which water = XYZ, or water = PQR, are equally close to the actual world as scenarios in which water = H<sub>2</sub>O, and there are more scenarios in which water fails to be H<sub>2</sub>O than in which water is H<sub>2</sub>O.

To get this to fit with the letter of Safety<sub>1</sub>, we need to make explicit one feature of the measure of 'distance' between worlds. If we opt for Sensitivity<sub>1</sub>, we can understand 'distances' between scenarios in the standard way Lewis (1979) understands distances between possible worlds. According to Lewis, differences in laws of nature are responsible for larger distances between worlds than differences in particular matters of fact, and, naturally, the larger the differences of either law of particular fact, the larger the distances between worlds. If we opt for Safety<sub>1</sub>, we are obliged to say something about how close complete scenarios are to partial scenarios. All things being equal, scenarios in which water = XYZ, or water = PQR, are *equally close* to the actual world as scenarios in which water = H<sub>2</sub>O, but this does not yet tell us that any of them are *close*. To get the desired result, we have to differentiate between epistemically possible scenarios that *add* to the partial scenario *w'* which verifies Joe's belief-forming method at the actual world and epistemically possible scenarios that *diverge* from facts verified by *w'*. Recall that *w'* is 'agnostic' on any number of issues. Adding to *w'* involves filling in the gaps in *w'*, whereas diverging from *w'* involves contradicting *w'* on something about which *w'* is *not* agnostic. A world *w* *adds* to *w'* if some facts that are verified by *w* are not verified by *w'*. A world *w* *diverges* from *w'* if some facts that are verified by *w'* are not verified by *w*. 'Differences' between *w'* and *w* that are mere additions, do not add distance between the scenarios, but 'differences' between *w'* and *w* that are divergences do. In short, the differences that add *distance* are not additions to the partial scenario, but divergences from it.<sup>23</sup> This gives us the result that there are many close possible scenarios in which water = XYZ, or water = PQR (and so on), but in which Joe

<sup>23</sup> A bit more formally: a world *w* diverges from the partial scenario *w'* which verifies  $\{\psi : \psi \in \Sigma(S, @)\}$  if and only if there is a sentence *s* of  $\mathcal{L}$  such that *s* is verified by *w* and  $\neg s \in \{\psi : \psi \in \Sigma(S, @)\}$ . A world *w* adds to the partial scenario *w'* which verifies  $\{\psi : \psi \in \Sigma(S, @)\}$  if and only if there is a sentence *s* of  $\mathcal{L}$  such that *s* is verified by *w* and  $s \notin \{\psi : \psi \in \Sigma(S, @)\}$ .

continues to form his belief about the target proposition in the same way as in the actual world and in which his belief is false. Joe's belief is not  $\text{Safe}_1$ .

Contrast this with good cases. Suppose you base your beliefs about the chemical composition of water on a good textbook. That water is  $\text{H}_2\text{O}$  is not something we can know a priori, so there are 1-possible worlds where the chemical composition of water is something else, XYZ say. However, the content of the textbook is informed by the work of people whose practices actually *track* the chemical composition of water. In scenarios in which the chemical composition of water is not  $\text{H}_2\text{O}$ , the textbooks will reflect this fact. As such your belief is  $\text{sensitive}_1$ : in the closest scenario(s) in which your belief that water is  $\text{H}_2\text{O}$  is false, you do not come to believe that water is  $\text{H}_2\text{O}$  via the method you actually use in forming the belief that water is  $\text{H}_2\text{O}$ . That the textbook is informed by the work of people whose practices track the chemical composition of water also ensures that your belief is  $\text{Safe}_1$ . Because your belief-forming method tracks the chemical composition of water across possible scenarios, it follows that in all of the nearby and very nearby scenarios in which you continue to form your belief about the target proposition in the same way as the actual world, your belief continues to be true.

$\text{Safety}_1$  and  $\text{Sensitivity}_1$  also deal with another standard case of epistemic luck concerning necessary truths:

### *Mathema*

Mathema uses a calculator to find out the sum of  $12 \times 13$ . As a result, he forms a true belief that  $12 \times 13 = 156$ . Unbeknownst to Mathema, however, his calculator is in fact broken and generating "answers" randomly. (Pritchard 2012, 256)

It is easy to see that Mathema's belief is  $\text{Unsafe}_1$ . Since the calculator is generating answers randomly, there are many close scenarios in which Mathema continues to form her belief about the target proposition in the same way as the actual world, but where the calculator generates the wrong answer. Mathema's belief is also not  $\text{Sensitive}_1$ . Though there can be many different ways to compute a function, in working calculators, it is true to say that for any numbers  $x$ ,  $y$ , and  $z$ , such that  $x \times y = z$ , the calculator will produce the output ' $z$ ' given the input ' $x \times y$ ' *because*  $x \times y = z$ . Properly functioning calculators reliably tell us the mathematical facts, whatever those mathematical facts *might* (from the point of view of one's epistemic predicament) be. Properly functioning calculators *track* the mathematical facts across epistemically possible scenarios. In the closest scenario in which  $12 \times 13 \neq 156$ , a properly functioning calculator will give the right result. This, however, is clearly not the case with calculators that generate answers randomly. Recall that Becker individuates methods narrowly and in a content-specific way. The belief-forming method here would be *If I see the calculator display '12 × 13 = 156', then 12 × 13 = 156*. However, were the proposition that ' $12 \times 13 = 156$ ' false, Mathema would continue to form the belief that ' $12 \times 13 = 156$ ' via the method she actually uses in forming that belief. Mathema's belief is not  $\text{Sensitive}_1$ .

What about someone who learns mathematical truths through testimony? So long as each link in the testimonial chain, including the first, is  $\text{Safe}_1/\text{Sensitive}_1$ , then the person who receives the testimony can gain  $\text{Safe}_1/\text{Sensitive}_1$  mathematical beliefs, so long as their own belief-forming method is  $\text{Safe}_1/\text{Sensitive}_1$ . On the other hand, if, for

instance, the testifier is Gullible Joe, who forms his mathematical beliefs on the basis of looking at his lucky 8-ball, then Joe's epistemic luck will be transmitted to the person receiving the testimony.<sup>24</sup>

Relatedly, one happy result of Safety<sub>1</sub> is that it deals nicely with a class of counterexamples to the analysis of epistemic luck in terms of Safety<sub>2</sub>, suggested by Lackey:

[C]onsider [a person, Penelope,] winning through a lucky guess a game show that presents contestants with multiple choice options. Now imagine that there is a feature,  $\phi$ , of the final winning answer that is entirely disconnected from its correctness but is such that its presence will invariably lead to Penelope to choose that answer. Suppose further that the current producer of the show, Gustaf, has a similar obsession with  $\phi$ , so that he ensures that the final winning answer of the day will possess this feature. Perhaps  $\phi$  is being presented in the color purple, so that when in doubt Penelope will invariably choose the answer displayed in purple and Gustaf will always present the final winning answer in purple. (Lackey (2008, 263))

Despite being a paradigmatically lucky event, Penelope's guess is safe<sub>2</sub>, as there are no close possible worlds in which Gustaf does not present the winning answer in purple and no close possible worlds in which Penelope picks an answer which is not presented in purple. The fact of Gustaf's purple fixation '*just happen[s] to fortuitously combine*' with Penelope's similar obsession, to make the event safe<sub>2</sub>. Lackey suggests a recipe for constructing such counterexamples:

[F]irst choose a paradigmatic instance of luck, such as winning a game show through a purely lucky guess, emerging unharmed from an otherwise fatal accident through no special assistance, etc. Second, construct a case in which, though both central aspects of the event are counterfactually robust, there is no deliberate or otherwise relevant connection between them. Third, if there are any residual doubts that such an event [is Safe<sub>2</sub>] add further features to guarantee counterfactual robustness across nearby possible worlds. (ibid.)

Here, Safety<sub>1</sub> can be put to use. Recall that Safety<sub>1</sub> orders scenarios relative to the agent's belief-forming method— $\{\psi : \psi \in \Sigma(S, @)\}$ —at the actual world. Gustaf's obsession with purple is not part of Penelope's basis for the belief that the correct answer will be presented in purple. Hence, there are many scenarios close to  $\{\psi : \psi \in \Sigma(S, @)\}$  in which Gustaf does

<sup>24</sup> I thank an anonymous referee for the suggestion that I apply the account to Mathema and the case of belief in necessary truths as a result of testimony. The referee also suggested that I apply the account to Mišćević's (2007, 49) 'Jane' case: 'Take the case of two mistakes in calculation or proof that cancel each other out resulting in the correct solution. Suppose they are extremely hard to detect, so that the thinker, call her Jane, is justified in trusting her calculation. Jane is Gettier lucky in her final belief. Cases of luck like the Jane example are only to be expected in matters of mathematical problem solving and theorem proving.' Here, I think, something else is going on: Jane is attempting to gain knowledge a priori. It seems to me however that Mišćević is wrong to ascribe justification to Jane. If Jane makes two errors in her calculation or proof, then she cannot have understood that calculation or proof. In failing to soundly follow the correct rules of deductive inference Jane is attempting to make use of, Jane lacks the kind of a priori justification needed for a priori knowledge.

not present the winning answer in purple and, as such, many nearby scenarios in which Penelope answers incorrectly. Safety<sub>1</sub> gives the correct result that picking the correct answer is lucky *for* Penelope, *given the way she forms her belief*.<sup>25</sup>

## 4 Applications

No doubt more could be said about the foregoing, but here the goal is simply to sketch out what sort of shape this research programme could take, rather than trying to settle in advance how every detail must be parsed. It is worth mentioning, again in suitably open and nascent form, two applications: one from the philosophy of religion and the other from the philosophy of mathematics. Take the problem of religious diversity. It is often thought that the diversity of religious traditions, along with people's tendency to adhere to the religious tradition of their birth creates an epistemological problem for religious belief. Hank was born in the USA and believes in a trinitarian God. Yet, if he was born in Indonesia, he would believe in a unitarian God. Observations such as these are the starting point for reflection on the epistemological problems of religious diversity.<sup>26</sup> We may wonder if Hank's religious beliefs, if true, are the product of epistemic luck. A natural and attractive way to think about these issues is in terms of truth-tracking. Hank, we might think, is tracking the beliefs of his religious community and not the existence of God; Hank's beliefs are sensitive to the contents of his religious tradition but not sensitive to the existence and nature of God. But if the existence and attributes of God are metaphysically necessary then, so long as religious believers lock onto their beliefs in a stable way, the sensitivity<sub>2</sub>, and the safety<sub>2</sub>, of their beliefs are guaranteed. The point here is not to insist the way to understand the epistemological problems of religious diversity in terms of a sensitivity condition, but rather that unless a modal analysis of knowledge can handle necessary truths, *any* such analysis will be of limited use in coming to shed light on the matter. This isn't to deny that much of interest and import can and has been said about religious diversity without appealing to modal epistemology, or even that traditional sensitivity<sub>2</sub> or safety<sub>2</sub> might have something to contribute, but that a full appreciation—let alone a resolution—of these issues is simply foreclosed until we have an understanding of epistemic luck that can be applied to them in all their fullness.

A perennial issue in the philosophy of mathematics is the set of epistemological problems surrounding mathematical objects. Mathematical objects are abstract, in the sense that they have no causal powers. This is often thought to generate an 'access problem' for mathematical objects; how can we have knowledge of the existence of mathematical objects if they, by their very nature, can make no difference to anything observable? The *locus classicus* for this sort of worry is Benacerraf (1973), but Benacerraf states the problem in terms of a causal theory of knowledge that is now widely rejected. Field (1989) has another shot at pinning it down, this time in terms of the challenge of explaining the reliability of our beliefs about mathematical objects. But the same sorts of problems arise. Since mathematical objects are usually understood as

<sup>25</sup> It should also be clear now how Safety<sub>1</sub> and Sensitivity<sub>1</sub> can handle the Parenthood Test case discussed above.

<sup>26</sup> See Basinger (2016 §3–4) and Pittard (2017) for an overview.



being necessarily existent, so long as believers in mathematical objects lock onto their beliefs in a stable way, the reliability (and sensitivity<sub>2</sub>, and safety<sub>2</sub>) of their beliefs is guaranteed. Again, a natural and attractive way to frame the problem is in terms of truth-tracking. Because mathematical objects are abstract, were they *not* to exist we would continue to believe that they do using the belief-forming methods we actually use. Sensitivity<sub>1</sub> captures this; Sensitivity<sub>2</sub> does not. Again, the point here is not to insist that understanding the problem in terms of Sensitivity<sub>1</sub> is obligatory, but that an appreciation of exactly what the epistemological problem with mathematical objects is, and whether it can be resolved may *require* some sort of modal analysis, and that any such modal analysis will have to involve epistemic rather than metaphysical possibilities.

Modal conditions, such as Safety and Sensitivity, have undergone a number of changes since then were first proposed. The core of modal epistemology, however, remains. I have already mentioned some refinements that are needed for anti-luck conditions parsed in terms of epistemic possibility, but it is likely that the back and forth process of testing against, and refining in the light of, various concrete thought experiments, would throw up some more. What I have suggested here is programmatic in the same way as Safety and Sensitivity always were, but the core—that a successful modal account of epistemic luck requires *some* such shift—is surely something we ought to accept.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Awodey, S. (2007). Carnap's quest for analyticity: The Studies in Semantics. In M. Friedman & R. Creath (Eds.), *The Cambridge Companion to Carnap* (Cambridge Companions to Philosophy, pp. 226–247). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CCOL9780521840156.011>
- Basinger, David. 2016. Religious Diversity (Pluralism). In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2016. Metaphysics Research Lab, Stanford University.
- Becker, K. (2007). *Epistemology modalized*. New York: Routledge.
- Benacerraf, Paul. 1973. Mathematical truth. *The Journal of Philosophy* 70 (19): 661–679.
- Black, T., & Murphy, P. (2007a). In defense of sensitivity. *Synthese*, 154(1), 53–71.
- Black, T., & Murphy, P. (2007b). In defense of sensitivity. *Synthese*, 154(1), 53–71. <https://doi.org/10.1007/s11229-005-8487-9>.
- Black, T., & Murphy, P. (2012). Sensitivity meets explanation: an improved counterfactual condition on knowledge. In K. Becker & T. Black (Eds.), *The sensitivity principle in epistemology* (pp. 28–42). New York: Cambridge University Press.
- Blome-Tillmann, M. (2017). Sensitivity actually. *Philosophy and Phenomenological Research*, 94, 606–625.
- Bolos, A., & Collin, J. H. (2018). A sensitive virtue epistemology. *Synthese*, 195(3), 1321–1335.
- Chalmers, David. 2002. Does conceivability entail possibility? In *Conceivability and Possibility*. Oxford University Press.
- Chalmers, D. (2006). *The foundations of two-dimensional semantics*. In *Two-dimensional semantics: Foundations and applications*. Oxford: Oxford University Press.
- Chalmers, David. 2011. The nature of epistemic space. In *Epistemic modality*, 60–107. Oxford University Press.
- Dretske, F. (1971). Conclusive reasons. *Australasian Journal of Philosophy*, 49, 1–22.
- Egan, A. (2011). In B. Weatherson (Ed.), *Epistemic modality*. Oxford: Oxford University Press.
- Field, Hartry. 1989. *Realism, mathematics and modality*. Wiley-Blackwell.
- García-Carpintero, M., & Macia, J. (2006). *Two-dimensional semantics*. Oxford: Oxford University Press.
- Goldman, A. (1976). Discrimination and perceptual knowledge. *The Journal of Philosophy*, 73(20), 771–791.



- Greco, J. (2009). *Achieving knowledge*. Cambridge: Cambridge University Press.
- Greco, J. (2012). A different virtue epistemology. *Philosophy and Phenomenological Research*, 85(1), 1–26.
- Greco, John. manuscript. Knowledge, virtue and safety. Man.
- Hales, S. D. (2016). Why every theory of luck is wrong. *Noûs*, 50, 490–508.
- Hirvelä, Jaakko. Forthcoming. Global safety: how to deal with necessary truths. *Synthese*.
- Lackey, J. (2008). What luck is not. *Australasian Journal of Philosophy*, 86(2), 255–267.
- Lewis, D. (1973). *Counterfactuals*. Oxford: Blackwell.
- Lewis, D. (1979). Counterfactual dependence and time's arrow. *Noûs*, 13, 455–476.
- Lewis, D. (1986). A Subjectivist's guide to objective chance, in his philosophical papers. Oxford: OUP, ii. 83–113.
- Luper, S. (1984). The epistemic predicament: knowledge, Nozickian tracking, and scepticism. *Australasian Journal of Philosophy*, 62(1), 26–49.
- Melchior, Guido (2017). Epistemic luck and logical necessities: armchair luck revisited. In Smiljana Gartner Bojan Borstner (ed.), *Thought Experiments between Nature and Society. A Festschrift for Nenad Mišćević*. Cambridge: Cambridge Scholars Publishing. pp. 137–150.
- Mišćević, N. (2007). Armchair luck: Apriority, intellection and epistemic luck. *Acta Analytica*, 22(1), 48–73. <https://doi.org/10.1007/BF02866210>.
- Nozick, R. (1981). *Philosophical explanations*. Cambridge: Cambridge University Press.
- Pittard, John. 2017. Religious disagreement. Internet Encyclopedia of Philosophy.
- Pritchard, D. (2005). *Epistemic luck*. New York: Oxford University Press.
- Pritchard, D. (2007). Anti-luck epistemology. *Synthese*, 158(3), 277–297. <https://doi.org/10.1007/s11229-006-9039-7>.
- Pritchard, Duncan. 2008. Knowledge, luck and lotteries. In *New Waves in Epistemology*, edited by Vincent Hendricks. Palgrave Macmillan.
- Pritchard, D. (2009). Safety-based epistemology: whither now? *Journal of Philosophical Research*, 34, 33–45.
- Pritchard, D. (2012). Anti-luck virtue epistemology. *Journal of Philosophy*, 109, 247–279.
- Roland, J., & Cogburn, J. (2011). Anti-luck epistemologies and necessary truths. *Philosophia*, 39, 547–561.
- Roush, Sherrilyn. 2005. *Tracking truth: knowledge, evidence, and science*. Oxford University Press.
- Russell, B. (1948). *Human knowledge*. London: Allen & Unwin.
- Sainsbury, R. M. (1997). Easy possibilities. *Philosophy and Phenomenological Research*, 57(4), 907–919.
- Sosa, E. (1999). How to defeat opposition to Moore. *Noûs*, 33(13), 141–153.
- Sosa, E. (2007). *A virtue epistemology: apt belief and reflective knowledge, volume I*. Oxford: Oxford University Press.
- Sosa, E. (2009). *Reflective belief*. Oxford: Oxford University Press.
- Sosa, E. (2010). How competence matters in epistemology. *Philosophical Perspectives*, 24(1), 465–475.
- Stalnaker, R. (1968). A theory of conditionals. In N. Rescher (Ed.), *Studies in logical theory* (pp. 98–112). Oxford: Blackwell.
- Weatherston, B. (2004). Luminous margins. *Australasian Journal of Philosophy*, 82(3), 373–383.
- Williamson, T. (2000). *Knowledge and its limits*. Oxford: Oxford University Press.
- Williamson, T. (2009). Replies to critics. In Duncan Pritchard & Patrick Greenough (eds.), *Williamson on Knowledge*. Oxford: Oxford University Press. pp. 279–384.
- Zagzebski, L. (1996). *Virtues of the mind: an inquiry into the nature of virtue and the ethical foundations of knowledge*. Cambridge: Cambridge University Press.
- Zagzebski, L. (1999). What is knowledge? In J. Greco & E. Sosa (Eds.), *The Blackwell guide to epistemology*. Oxford: Blackwell.